Session 5: Income and Wealth: Theory and Practice I
Time: Thursday August 9, 2012 AM

Paper Prepared for the 32nd General Conference of
The International Association for Research in Income and Wealth

Boston, USA, August 5-11, 2012

Well-Being and Measurement Error of Income Reported in a Social Survey versus Income Recorded by a Tax Administration

Dmitri Romanov and Yury Gubman

For additional information please contact:

Name: Dmitri Romanov
Affiliation: Israel Central Bureau of Statistics

Email Address: dromanov@cbs.gov.il

Well-Being and Measurement Error of Income Reported in a Social Survey versus Income Recorded by a Tax Administration

Dmitri Romanov* and Yury Gubman**

Abstract

In this study, measurement error in the reportage of individuals' gross labor income by means of a banded question is analyzed by matching responses in Israel's Social Survey to the respondents' records in the income tax administration. We find that regression to the mean occurs in the income reported in the survey as against that obtained from the administrative file. Positive measurement errors are more common among low-income respondents and negative measurement errors are more common among those of high income. Consequently, Gini index based on income reported in the survey is by 10% lower than that measured on the basis of the same individuals' income in the administrative file. For employees, we find a positive effect of unstable employment and fringe benefits on overstating the individual's labor income in the survey. For the self-employed, few of the economic and demographic variables were useful in explaining the measurement error. The analysis proves that pooling employees and the self-employed, high-income and low-income respondents, and positive and negative measurement errors in one model may seriously bias the estimates of various factors that are unique to each group.

1. Introduction

Income is a crucial variable in profiling the socioeconomic status of an individual, a household, or a geographical region in statistical publications and in the copious research that has been done in the social sciences at large. Analysis of income distribution and income inequality is a research topic in itself, one closely related to poverty levels and economic disparities in a society. Therefore, one cannot overstate the gravity of the implications of measurement errors in income variables that may

* Israel Central Bureau of Statistics, corresponding author, email dromanov@cbs.gov.il.

** Israel Central Bureau of Statistics and Hebrew University of Jerusalem.

introduce significant biases in empirical research findings and the policies flowing from them (Davern et al., 2005; Bound et al., 2001; Duncan and Hill, 1985).

Generally speaking, one may obtain income data at the individual level from two sources: surveys and data from administrative sources such as tax authorities or social security authorities, e.g., the National Insurance Institute in Israel. Survey data are prone to errors of sampling, measurement, and non-response, and are costly to gather. Individual-level administrative data are susceptible to biased reportage due to the natural relationship between reporting income and paying taxes or receiving benefits. While such sources do provide "census"-type coverage of a specific type of income, the data become available for statistical processing at a lag due to the time it takes the competent authority to gather them. Administrative sources seldom provide information about income that is not taxable or under the reporting threshold. Income data from both types of sources, administrative and survey, also differ in definitions, periods of reference, and populations covered (Yitzhaki, 2007).

Unlike the other surveys performed by the Israel Central Bureau of Statistics (CBS), which gather detailed information about individual and household income by types of income in order to analyze the income per se—detailed income for substantive analysis—persons enumerated in the Social Survey are asked to report labor income (from wage/salary and self-employment) only, placing it in one of ten predetermined income bands of their choice. This corresponds to the purposes of the survey, a survey of individuals in which labor income serves as a socioeconomic classificatory variable. Use of a banded question also makes it easier to respond and raises the rate of response to this item (Collins and White, 1996) in view of the known sensitivity of questions relating to income (Tourangeau and Smith, 1994; Groves, 1989). The band method is widely used for the reportage of income data in non-

economic surveys, e.g., those relating to health, education, and welfare issues and attitudes (Office for National Statistics, 2010); it is also employed in economic surveys that suffer from high non-response rates (Juster and Smith, 1998; Foster and Lound, 1993). Reconstruction strategies by respondents who lack clear or documented knowledge about their income level often make even "accurate" reporting approximate and rounded (Moore et al., 2001), making banded questioning all the more justified (Romanov and Furman, 2011; Czajka and Denmead, 2008).

Micklewright and Schnepf (2007) contend that the use of a banded question may result in loss of information about income distribution within the bands and, in turn, may impair the accuracy of the aggregate estimates derived from the survey. The authors compare income distribution in surveys that ask a banded question—relating to total personal/household income from all workplaces—with that elicited by surveys that feature detailed reporting, as a continuous variable, according to type of income (which also suffer from reporting problems). Micklewright and Schnepf conclude that personal income reported in a banded question is more accurate than total household income and is more accurate among men than among women. Most of the disparities found relative to detailed-item surveys, they say, focus on the lower end of the income distribution and are associated with a large proportion of total income originating in transfer payments and social support, which are notorious for faulty reportage in surveys due to under-coverage of benefits and social support and under-reporting of these types of income. These findings are consistent with those of Mathiowetz et al. (2002) and Moore et al. (2001).

In their edifying review, Moore et al. (2001) present the abundant findings of twenty years of methodological and empirical research into the reliability, biases, and reporting errors about income of various kinds (wage labor, social benefits, property).

When income estimates from important American surveys (SIPP and CPS) are compared with administrative sources, it is found that:

> Wage/salary income source reports may suffer a very modest level of net under-reporting bias; response bias estimates for wage/salary income amounts are generally small and without a consistent sign, indicating neither under- nor over-reporting bias given accurate source reporting (p. 356).

Moore et al. do not relate to problems in the reporting of income from business and self-employed activity, which are more difficult in terms of conceptualization, definition, questioning, and reporting than wage/salary income and transfer payments (Dror-Cohen, 2008; Martin et al., 1996); their measurement poses a real challenge to those in charge of the surveys.

Another issue that usually eludes the attention of researchers who focus on investigating reportage for measurement errors and summarize their findings as matters of net response (or measurement) bias is the distribution of the errors. If, for example, errors in income reportage in a survey correlate negatively with “true” income (Bound and Krueger, 1991; Romanov and Furman, 2006) so that errors among the wealthy in one direction are offset by errors among the poor in the opposite direction, the average error may be negligible but the estimates of inequality, polarization, and indicators of poverty would be biased. The research literature, however, associates measurement errors at the low end of the income distribution with under-reportage of benefits and support (Hansen and Kneale, 2011; Meyer and Sullivan, 2003) and finds rampant under-reporting of capital and property income at the upper end (Micklewright and Schnepf, 2007; Moore et al., 2001). Ostensibly,

then, measurement errors are in the same direction, i.e., under-reporting exists at both extremes.

As for the factors that should explain the errors in measuring personal labor income—the variable at the focus of this study—the literature alludes to occupational mobility, the frequency and duration of spells of joblessness, multiple posts/employers and their characteristics, the form of remuneration and the period for which is received, and variance in hours worked during the reporting period of the survey (Abowd and Stinson, 2011; Akee, 2011; Mathiowetz et al., 2002; Rodgers et al., 1993).

Accordingly, the current study has two main goals. First, it aims to compare banded labor income reported in a social survey with information from an administrative file (that of the Israel Tax Authority), matched at the individual level, and to estimate disparities between the labor income distribution reported in the survey as against that reported in the administrative source, distinguishing among differences at the extremes of the distribution that have a sizable effect on poverty and inequality estimates and separating employees' wage/salary income from business income of the self-employed. The second goal is to analyze factors that may explain the disparities between the two sources in individuals' reported income. Here reference to three types of factors will be made: those related to question design and the cognitive process of response; those related to differences between employees and the self-employed in the nature and definition of labor income; and employment-related factors that make it hard to correctly reconstruct and report the level of labor income in the survey, even in a banded question.

The rest of the article is organized as follows: Part 2 presents the data and analyzes relations among the main investigation variables by means of the NLMA

curve, a new graphic tool for the analysis of monotonic relations. Part 3 shows the results of the estimation and Part 4 concludes.

2. Research Data and Variables

2.1 Database and Definition of Variables

The study is based on data from the 2008 Social Survey, performed each year in Israel by the Central Bureau of Statistics. In 2008, a systematic random sample of 8,899 individuals who belong to the survey population was extracted from the Population Registry; all were residents of Israel aged 20+. Among them, 7,327 responded (response rate of 82% of eligible persons). The survey was conducted by means of a computer-assisted personal interview in the sampled person's home.

One of the main characteristics of the survey population is its level of labor income. This variable was constructed on the basis of a personal question about the respondent's gross income from all sources (hereinafter: "gross income"). The question was asked after employment was investigated and labor status was defined: employee, self-employed, manager of owned company, or member of cooperative. The wording of the question for employees was: "Last month, what was your gross income, before deductions, from all places where you worked?" Self-employed respondents and those who managed their own companies were asked: "Last month, what was your gross income, before deductions, from all places where you worked, including wages and income from a business?" The question was administered only to individuals who reported working as employees or self-employed in the month preceding the survey. Of the 4,493 survey participants who were asked the question,

8.26% did not respond (refused or did not know what to say),[1] lowering the number of respondents to the gross income item to 4,122. To answer the gross income question, the respondent was asked by the interviewer to avail him/herself of a card on which ten bands, specified in Table 1, were listed.

The gross income question was immediately followed by one about net income: "What was your net income after deductions such as income tax, social security and mandatory health insurance contributions?" The question was added to the survey in 2007, in response to findings of Romanov and Furman (2006), as a tool with which respondents could control the distinction between gross income and net income by themselves. When respondents are asked about their net income after having answered the gross income question, if they mistakenly cited their net income in response to the first question, they are expected to realize their mistake and go back and correct the mistaken gross income variable.

In 2008, the topic of job mobility was examined at length in the Social Survey. In this context, the survey participants were asked about their history in the labor market, the components of their wages, their fringe benefits, and their attitudes toward and expectations about their present and future status as employees/self-employed. We used these data in our analysis as adjuncts to the respondents' socioeconomic and demographic indicators, thereby enriching the analysis with information that general surveys usually lack.

[1] The item non-response rate to the gross income question in Israel's Social Survey is low by the standards of main U.S. government surveys, in which rates reach 25% or more (Moore et al., 2001; Czajka and Denmead, 2008; Groves et al., 2001). Gubman and Romanov (2010), addressing the sensitivity of non-response to the gross income question, found that it does not sustain the standard missing-at-random assumption.

The Social Survey is performed by means of Blaise, which, in a special application known as Audit Trail (AT), allows for the creation of a log file that records every motion the interviewer does during the interview. This log captures the following variables among others: interviewer number, respondent number, time of entrance to fill in a field (a question), time of exit from field, value of variable upon entering field, and value entered by interviewer. If the respondent corrects a previous response, as many lines are recorded in the relevant field in the AT file as the number of times the interviewer repeated the specific question and corrected the value of the field.

Basing ourselves on the AT file, we calculated three variables for each respondent who answered the gross income question: Did the respondent go back to the question? (0 if s/he answered once and 1 if s/he went back to the question and corrected h/her previous response); time (in seconds) that it took the respondent to answer the gross income question; and the size of the correction to gross income when the question was repeated, calculated as the difference between the first answer and the last one.

In addition, to verify the income reported in the survey against actual income, we used information obtained from the income tax authorities. The individual-level records were linked with various sources of information on the basis of a unique ID number. The income tax data for employees itemize all jobs that the individual held during the tax year, annual wage/salary and the months in which s/he held them in each job. We calculated the average monthly income of any employee in any job by dividing the annual income by the number of months in which the job was held. To match the Social Survey's question of gross income from all jobs, employees' income in the month preceding the enumeration month from the income tax file was

calculated as the total average gross income from all posts that the individual held in the pre-enumeration month. Since the self-employed do not report to the tax authorities working months, we had no information on whether an individual worked in a given month or not. Consequently, we calculated the gross monthly income of the self-employed by dividing the reported annual income by 12.

Individuals for whom records were not found in the income tax file or for whom the AT file contained no valid records (usually for technical reasons) were deleted from the database. The final number of observations for which the three sources of information (social survey data, income tax data, and an AT file) provided data was 3,417. The final database comprised records of 408 self-employed (12%) and 3,009 employees (88%). Table 1 shows the gross income values and the distribution of the responses obtained in the survey.[2]

Figure 1 contrasts the income reported in the survey with the same individuals' income in the administrative file, grouped into the same bands as are specified in the survey. The thick line in the graph denotes income from the administrative source within the survey bands. The thin line is the mean of the responses in the survey that were given by the same individuals who were placed within this band in accordance with their income as shown in the administrative source. The broken line plots the trend of the mean responses in the survey. The X-axis shows centiles of income from the administrative source.

Figure 1 indicates that low-income respondents (bands 1, 2, and 3) over-reported their labor income in the survey relative to the administrative source and that those of relatively high income (bands 5 and above) under-reported it. This phenomenon—

[2] All analyses in this study were performed on the basis of the final calibrated weights that CBS produced for the 2008 Social Survey.

over-reporting of low income in the survey and under-reporting of high income—is known in statistical terms as regression to the mean.

Table 2 presents indices of the distribution of the specified variables on the basis of data in the aforementioned AT file. We see that 22% of the respondents went back to the question at least once and that most corrections were made in an upward direction, by about half a band on the scale of responses in the survey. By conjecture, they made this correction after realizing, when asked the net income question, that when responding to the previous question they had stated a net income value instead of a gross income value. Then, realizing this, they returned to the gross income question and offered a higher value. By implication, adding a net income item to the questionnaire improves the accuracy of reportage about gross income. It also allows those who do not know their gross income to respond. (In this study, those who did not answer the gross income question were omitted.)

It was also found that, on average, self-employed participants took much longer to respond than employees did (27.7 seconds as against 21.8, respectively) and that their standard deviation was greater as well (22.3 vs. 17.8).

2.2 Defining the measurement error variable

Most studies concerning measurement error treat income data from the administrative source as "true". This approach assumes that, ideally, survey respondents would report their income exactly as it appears in the tax authorities' files. However, Moore et al. (2001) disapprove of this approach, emphasizing:

> Data from independent sources are almost never completely comparable to the survey data—due to sampling frame differences, timing differences, definitional differences, etc.—and the adjustments necessary to make them comparable are often inadequate. The flawed

> adjustments, and the fact that the independent estimates themselves are subject to various errors and omissions, add uncertainty to any comparison of survey and benchmark estimates (p. 333).

Some of these constraints fall away when survey reportage in the survey is compared with information from an administrative source at the individual level, as is done in the present study. However, discrepancies between these sources in the definition of income and the timing of its recording definitely exist; they affect the inferences that may be drawn from such a comparison.[3]

In contrast to the common practice, Abowd and Stinson (2011) assume as a point of departure that administrative data are not a "gold standard" due to errors in employers' reportage to the tax authorities and the existence of black-market employment and income. Furthermore, any comparison of administrative data and survey reportage at the individual level entails record linkage, which in itself may bring about errors.

Accordingly, we treated discrepancies discovered in a comparison between the survey reportage and administrative-source data not as a "reporting error" by the individual but as a "measurement error" that is affected by the full range of factors.

When the survey data are categorical, a "measurement error" variable may be defined in several ways. In particular, one may calculate the expectation in each of the bands specified in the survey. Let F represent the distribution of income from the administrative source (as a continuous variable) across the population of survey

[3] Notably, in Israel, unlike the U.S., self-reportage to the tax authorities is compulsory only for the self-employed and for managers of companies that they own. Employees' wage/salary income is reported to the tax authorities by employers, who withhold income tax, social security and mandatory health insurance contributions at source.

respondents. For each labor-income band specified in the survey, an estimator of the expected income from the administrative file that would have been reported in the same band had it not been for measurement errors, according to distribution *F*, will be:

$$(1)\ \hat{\mu}_k^F = n_k^{-1} \sum_{i \in G_k} \hat{y}_i$$

where:

$\hat{y}_i$ —the predicted value of gross administrative income of individual *i* from all jobs, calculated on the basis of estimated distribution $\hat{F}$ in income band G_k as defined in the survey.

n_k —the number of individuals in group k.

To find an estimator for distribution *F*, we fitted by maximum likelihood several accepted parametric distributions to the continuous administrative data of the survey respondents' gross income. We examined log-normal and Gamma distributions (Banerjee et al., 2006), elliptic, asymmetric, skew-normal, and skew-t distributions (Azzalini and Capitanio, 2003). The best fit, i.e., the one that minimizes the deviations of the actual data from the value predicted on the basis of a theoretical parametric distribution, was obtained for the skew-t distribution. Figure 2 presents the distribution of the survey respondents' administrative income and the estimated distribution line.

Having estimated distribution function *F* under the assumption $y_i \sim \hat{F}$, we may calculate the measurement error of individual *i* (in%) by using Formula (2):

$$(2)\ e_i = \left(\frac{\hat{\mu}_k^F - y_i}{\hat{\mu}_k^F}\right) * I(i \in G_k) * 100$$

where I is an indicator that is assigned the value of 1 if the "true" income of individual *i*, as shown in the administrative file, does not belong to the band in which the individual reported h/her income in the survey; otherwise, it is 0.

The measurement error calculated by Formula (2) is positive when the income recorded in the administrative file is smaller than the expectation of the response in the survey (a case of over-reporting in the survey) and negative in the opposite case (under-reporting in the survey).

Table 3 presents main indicators of the distribution of measurement error by respondent's labor status. It may be seen that, on average, employees are more inclined than the self-employed to negative measurement errors (44.6% vs. 24%, respectively). The share of respondents who have no measurement error is higher among employees (40.4% vs. 33.8%, respectively) and the variance of measurement errors is wider among the self-employed (S.D. 97.7 vs. 67.3). By implication, there are meaningful differences between employees and the self-employed in the distribution of measurement errors.

2.3 Monotonic and Non-Monotonic Relations—Analysis of NLMA Curves

To examine the form of the relation between the main explanatory variables in this study and the explained variable—measurement error in gross labor income—we used the NLMA (Normalized Line of Independence minus Absolute concentration) graphic tool suggested by Yitzhaki and Schechtman (2012). This technique allows us to determine whether the relation between the explained variable and the explanatory variable is monotonic across the range of data—an inquiry that is needed to choose the functional form of the regression.

An NLMA curve starts at point (0,0) and ends at point (1,0). The horizontal axis shows the cumulative distribution of the independent variable, X, while the vertical axis presents the cumulative value of the dependent variable, Y. The line of independence is the line between (0,0) and (μ_Y,1), which may be interpreted as the absolute concentration curve of *Y* if *Y* were statistically independent of *X*.

The characteristics of the NLMA curve are such that, at segments in which the curve is concave (convex), the sign of OLS regression coefficient *Y* on *X* will be positive (negative) and when the curve is a straight line, the regression slope will be zero. If the NLMA curve does not cross the horizontal axis, the sign of the OLS regression coefficient will be unchangeable by any monotonic transformation of *X* and/or *Y*.

An analysis facilitated by NLMA can be performed only on non-binary variables. Therefore, the analysis that follows concerns three main explanatory variables: gross labor income (for the month preceding enumeration in the survey) from the administrative source, (first) response time to the gross income question, and size of correction to gross income value by those who went back to the gross income question during the interview.

Figure 3 shows NLMA curves that examine the monotonicity of the relation between the measurement error and gross income as reported in the administrative source for the full sample, employees, and the self-employed. It may be seen that the three curves obtained are convex across most of the distribution band. This implies the existence of a negative monotonic relation between income from the administrative source and measurement error in almost all of the distribution band, except the segments up to the 10^{th} percentile and over the 95^{th} percentile for the full sample and for employees, and the segment up to the 15^{th} percentile for the self-

employed. The regression coefficient of the administrative-income variable is expected to be negative and robust to monotonic transformations in this variable; therefore, we shall use the income from the administrative source in a natural logarithm.

Figure 4 shows NLMA curves for the relation between measurement error and response time to the gross income question for the full sample, employees, and the self-employed separately. Observing the full sample, we find an interesting pattern in the relation: among those who responded immediately—in the band that captures the first 10% of response-time values—the responses were generally accurate. The more time it took to respond, the greater the measurement error in the next 40% of the distribution; afterward, the errors contracted again but remained positive. Those who took an especially long time, in the upper 20% of the distribution, gave less-accurate responses. By conjecture, respondents in this band reconstructed data that varied widely or performed a complex calculation of their income. Looking at the employee and self-employed curves, we found a clear positive relation for employees in roughly the first half of the response-time distribution; afterwards, the relation was positive but unstable. Among the self-employed, in turn, nowhere in the distribution did we find a clear monotonic relation between response time and measurement error.

Figure 5 shows NLMA curves for the relation between measurement error and change in the value of the response to the gross income question during the interview—for the full sample, employees, and the self-employed. It may be seen that in the sample at large and for employees, within the band bounded by the 20^{th} and the 60^{th} distribution percentiles, there is a positive non-monotonic relation between the change in the respondent's answer as a result of h/her returning to the gross income question during the interview and the measurement error. For the self-employed, the

form of the relation between the variables cannot be defined. By implication, we identified a monotonic relation in the segments of distribution of the AT variables and measurement error among employees but not among the self-employed. Therefore, we may expect the AT variables to be insignificant in the regression for the self-employed.

The full-sample and employee curves are similar because employees accounted for 88% of the sample. It is also evident that in the "response time" and "change of response" variables, the positive correlation is observed up to approximately the 60^{th} percentile, indicating that one cannot expect to arrive at a significant estimate for those who have a positive measurement error and who populate the upper percentiles of the distribution; the same may be said about the self-employed collectively.

These findings give us a statistical justification for dividing the sample into four sub-populations: by labor status (employees and self-employed) and by sign of measurement error. We circumscribe this by noting that the analysis by means of NLMA is bivariate, while additional effects may play important roles in explaining the phenomenon investigated.

3. Analysis of Factors for Measurement Error

3.1 Econometric Model

To identify the contribution of the factors that affect the measurement errors and estimate the intensity of their effect, we estimate the following regression (in matrix notation):

$$(3)\ e = \alpha + \beta X + \gamma L + \delta AT + \kappa WB + \varepsilon$$

where:

e—measurement error;

X—set of individual's background variables ;

L—set of variables representing the individual's employment and wage terms and variables of h/her subjective assessment of job mobility;

AT—set of Audit Trail variables, constructed in the manner described above;

WB—well-being variables associated with the individual's economic situation: satisfaction with work, with income, and with household's economic situation;

ε —vector of residuals.

As for issues related to the use of subjective evaluation variables as explanatory variables, we base ourselves on Bertrand and Mullainathan's (2001) conclusions:

> [Subjective] data may be useful as explanatory variables. One must, however, take care in interpreting the results since the findings may not be causal. Subjective variables are in practice useful for explaining differences in behavior across individuals (p. 59).

The estimation method was the weighted OLS and the final weights of the 2008 Social Survey were used. To test for the OLS model assumptions, each model estimated was checked for residual independence, normality, and homoscedasticity. We also tested for multicollinearity in the model by analyzing a matrix of covariance of the estimated regression coefficients and a conditional number index that should not exceed the accepted threshold of 50.

3.2 Estimation Results

As noted in Part 2.3 above, the factors associated with measurement error were analyzed separately for employees and for the self-employed due to a material difference in the regularity of these populations' income and a conceptual difficulty in defining it among the self-employed. Also, the absence of monotonic relations

between the explained variable and variables representing the cognitive process of responding entails separate analysis of positive and negative measurement errors. Accordingly, for each group of respondents that was differentiated by labor status, we estimated Model (3) in three ways: in a sample of all observations, including a group of observations that had no measurement error; in a sample of observations that had a positive measurement error (over-reporting in the survey); and in a sample of observations that had a negative measurement error (under-reporting in the survey). Notably, the sign of the estimates depends on the sign of the error. For example, a positive estimate in the analysis of positive measurement errors is indicative of a factor that increases the error, whereas a positive estimate in an analysis of negative measurement errors signifies a factor that lowers the error (toward zero).

Table 4 presents the estimates for the full sample. First, the goodness of fit (adjusted R-square) is 0.05 in the model covering all observations and 0.05 and 0.49, respectively, in the separate models for negative and positive errors.[4] Consequently, a hypothesis that all explanatory factors have the same effect on errors in both directions is not supported by the data. Therefore, below we relate to models of positive and negative error separately, irrespective of the pooled model.

As was found in the NLMA analysis, the level of income obtained from the administrative file affects measurement error by lowering positive errors and increasing (in absolute terms) negative ones, with elasticity of roughly 0.2. Going back to the gross income question during the interview and correcting the response reduces negative errors, as we would expect if respondents offered a net income value in their first response instead of one relating to gross income. The estimate of the

[4] This is partly due to the inclusion of the 40% of observations that had no measurement error as this term is defined in Formula (2); in these observations, the explained variable has no variation.

response time was not significant in the models for any direction of error, as we would expect after having examined the relation through the NLMA technique. Other factors controlled for, men tended to report 7% larger positive errors and 11% smaller negative errors than women. Age affected measurement errors in both directions in a U shape. An increase in level of education was identified with an increase in positive error but had no effect whatsoever on negative error.

Being an employee as opposed to being self-employed had a downward effect on errors in both directions—18% on positive errors and 31% on negative ones. This finding is consistent with the general conclusion in Moore et al. (2001) that reporting errors are negligible, on average, in the reportage of wage income. Individuals who reported being afraid of losing their jobs exhibited a 7% larger negative error. Holding an additional job on the survey date increased positive error at a 5% rate. Finally, satisfaction with income is identified with a 2% increase in positive errors and a 7% decrease in negative errors.

Table 5, like the pooled model (employees and the self-employed together), shows that the goodness of fit of the employees-only model is very strong for cases of positive error (0.57) and relatively poor for cases of negative error and all observations (0.09 and 0.07, respectively). The findings concerning the effects of the aforementioned demographic and socioeconomic variables on the full sample are also valid for employees. The employees-only model is especially mindful of employment characteristics that are likely to affect measurement errors. Thus, the number of jobs held increases positive error (by 5%), as do overtime work (by 3%) and fringe benefits such as full pay for sick days (by 3%), contributions to advance-training fund (by 6%), company car (by 8%), coverage of transportation expenses (by 5%), and profit-sharing (by 7%). In contrast, employees who reported not having been

promoted in their current job tend to report their income with smaller positive errors by 11%. Other factors were found to have significant effects on negative reporting errors. First, part-time employment reduces negative error by 28%. For the given duration (full- or part-time) of individual's main job, the more hours worked in all jobs, the smaller the negative error is. The negative error tends to be smaller among employees who reported having received a raise in recent years at their current job (by 11%) and those who benefited from employer contributions to pension (by 21%) and reimbursement of transportation expenses (by 11%). A much larger negative error (68%) is typical of employees who believe they have not been promoted on the job.

Table 6 presents estimates for the self-employed only. In this model, too, goodness of fit is relatively strong for positive errors (0.32) and weak for negative ones (0.02). In this model, unlike the employees-only model, the respondents' demographic and socioeconomic indicators had no effect, with the sole exception of academic education at the baccalaureate level. The AT variables, too, were not found significant. The only factors that had statistically significant effects (with a disclaimer for the small number of observations) were subjective variables. Thus, satisfaction at work and the feeling that the respondent's standard of living has fallen in recent years are identified with an increase in positive error by 7%. Two variables related to expectations—that business turnover will increase and that the business will make progress generally—had a downward effect of 10% on positive errors.

4. Summary and Conclusions

This study analyzed errors in the measurement of labor income on the basis of a comparison of reportage in a social survey, by means of a banded question, with records in an administrative file obtained from income tax authorities. Following Abowd and Stinson (2011), we did not consider the contents of the administrative file

as true reportage against which the data reported in a survey should be examined and any difference should be regarded as a reporting error in the survey. We treated discrepancies between the administrative source and the survey as errors in the measurement of a latent variable. These measurement errors may be affected by a gamut of factors including differences in definitions, reporting periods, timing of the recording and receipt of income, non-response in the survey, under-reporting to the income tax authorities, reporting errors, errors of record linkage, and so on.

We focused on the relation between measurement error and three groups of factors: those associated with question design and the cognitive process of answering a question; those related to differences between employees and the self-employed in the nature and definition of labor income, and those associated with employment characteristics that make it difficult to correctly reconstruct and evaluate the level of labor income reported in the survey.

The income data from the two sources were compared by linking records at the individual level via a unique national ID number. Then, the reporting period was synchronized and source of income from the administrative file was adjusted to the labor status as reported in the survey, thereby unifying the reportage in the survey and the administrative file in terms of monthly labor income. To define a measurement error, we first estimated a latent continuous variable of income that would have been reported in the survey had the question not been defined in bands, on the basis of a theoretical distribution that was fit with data that were actually reported in bands. We then calculated the measurement error as the difference between the estimate of this latent variable and the same individual's income as reported in the administrative file. The measurement error was defined as zero if both values fell into the same income band that appeared in the survey, as negative if the individual's reportage in the

survey was lower than the administrative data, and as positive if the survey reportage exceeded the administrative data.

Our main finding is that regression to the mean occurs in the income reported in the survey as against that obtained from the administrative source. Positive measurement errors were more common among low-income respondents and negative measurement errors were more common among those of high income. Consequently, the indicators of variance and inequality based on income reported in the survey were lower than when measured on the basis of the same individuals' income in the administrative file: Gini index values of 0.4267 as against 0.4746.

Generally speaking, a significant negative monotonic relation was found between the income recorded in the administrative file and the measurement error. The relation was especially significant for employees and differed in intensity between positive and negative measurement errors (elasticity of 0.2 and 0.5, respectively). In other words, the higher the individual's income is according to the administrative source, the smaller h/her measurement error in the survey will be if the error is positive, and the greater it will be (in absolute terms) if the error is negative. This finding is tantamount to a mirror image of regression to the mean across the income distribution elicited by the administrative file.

According to our research hypothesis, high-income persons who report their income in a social survey tend to "forget" income that they received from additional jobs, overtime, self-employed income, fringe benefits, and nonrecurrent gains such as bonuses and profit-sharing. As a result, they tend to under-report their labor income in the survey relative to the administrative data. In contrast, low-income workers, who hold part-time and/or irregular jobs, tend in surveys to report the income they receive in a full month of work, a level that may be not representative of their average

income. Consequently, one expects them to over-report their income relative to the administrative data. The findings of our research confirm these hypotheses and, by so doing, prove that the factors related to negative measurement error are different from those that are associated with a positive measurement error.

Parsing the analysis of measurement errors by the respondents' labor status—employees vs. self-employed—we found that the two populations should not be pooled into one model due to material differences in the conceptual definition of income, how income is measured, volatility in income level during the year, and reporting on income in the survey as against reporting it to the tax authorities. The response time to the gross income question was 27% longer among the self-employed than among employees, indicating that the former found the question harder to answer. Furthermore, only a few factors apart from administrative income level were found to be related to measurement errors among the self-employed: subjective variables such as satisfaction with income and expectations of business progress in the near future. Notably, among the self-employed, positive errors were almost twice as frequent as negative ones but the average error was close to zero, with greater variance than in errors among employees. Among the latter, the average error was *minus* 22% and negative errors were three times as common as positive ones.

Analyzing the response process by means of Audit Trail variables, we found that inserting a question about net income immediately after inquiring about gross income, as a logical way to allow respondents to control their responses, caused 22% of the respondents to go back and check the accuracy of their responses to the gross income question. Reversion to the gross income question during the interview and correction of the answer reduced measurement errors among both employees and the self-employed but did so more among the former than among the latter.

As for the contribution of subjective variables associated with respondents' satisfaction with their work, their breadwinning, promotion on the job, and expectations, we found a significant relation between them and measurement errors. We infer from this, following Bertrand and Millainathan (2001), that these variables yield important information for the explanation of disparities among individuals in the size and direction of income-measurement errors. There is no doubt, however, that separate research is needed to understand the cognitive mechanism that translates individuals' outlooks and attitudes into measurement errors generally and survey reportage errors particularly.

In sum, we learned that the analysis of income measurement errors in a survey as against an administrative file, with no distinction made between employees and the self-employed, between high-income and low-income respondents, and among factors that are unique to each group, may bias the findings, diminish their explanatory power, and impair our understanding of ways to improve the data that illuminate this crucial variable.

References

Abowd, J.M., and Stinson, M.H. (2011). Estimating measurement error in SIPP annual job earnings: a comparison of Census Bureau survey and SSA administrative data. US Census Bureau Center for Economic Studies Paper No. CES-WP 11-20.

Akee, R. (2011). Errors in self-reported earnings: The role of previous earnings volatility and individual characteristics. J of Development Economics, vol. 96 (2), 409-421.

Azzalini, A., and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. Journal of the Royal Statistical Society: Series B, vol. 65, 367-389.

Banerjee, A., Yakovenko, V.M., and Di Matteo, T. (2006). A study of the personal income distribution in Australia. Physica A, vol. 370, 54-59.

Bertrand, M., and Mullainathan. S. (2001). Do people mean what they say? Implications for subjective survey data. The American Economic Review, vol. 91 (2), Papers and Proceedings of the 113 Annual Meeting of the American Economic Association, 67-72.

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. Handbook of Econometrics, vol. 5, Chapter 59, 3705-3843. Elsevier North Holland.

Bound, J., and Krueger, A. (1991). The extent of measurement error in longitudinal earnings data: do the two wrongs make a right? Journal of Labor Economics, vol. 9 (1), 1-24.

Collins, D., and White, A. (1996). In search of an income question for the 2001 Census. Survey Methodology Bulletin 39, 2-10.

Czajka, J.L., and Denmead, G. (2008). Income data for policy analysis: a comparative assessment of eight surveys. Mathematica Policy Research, Inc.

Davern, M., Rodin, H., Beebe, T.J., Call, K.T. (2005). The effect of income questions design in health surveys on family income, poverty and eligibility estimates. Health Services Research, vol. 40 (5p1), 1534-1552.

Dror-Cohen, S. (2008). Enquiry of the self-employed and their income in surveys. Central Bureau of Statistics Working Paper 34, Jerusalem (in Hebrew).

Duncan, G.J., and Hill, D.H. (1985). An investigation of the extent and consequences of measurement error in labor-economic survey data. Journal of Labor Economics, vol. 3 (4), 508-532.

Foster, K., and Lound, C. (1993). A comparison of questions for classifying income. Survey Methodology Bulletin 32, 1-7.

Groves, R.M. (1989). Survey errors and survey costs. John Wiley and Sons, Inc.

Groves, R.M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (2001). Survey nonresponse. John Wiley and Sons, Inc.

Gubman, Y., and Romanov, D. (2009). Nonparametric estimation of non-response distribution in the Israeli social survey. Paper presented at the American Statistical Association Annual Joint Meeting, Washington DC.

Hansen, K., and Kneale, D. (2011). Does how you measure income make a difference to measuring poverty? Centre for Longitudinal Studies, CLS Cohort Studies, WP 2011/1.

Juster, F., and Smith, J. (1998). Enhancing the quality of data on income and wealth: recent developments in survey methodology. Paper presented for the 25 general Conference of the IARIW, Cambridge, England.

Martin, J., Cheesebrough, S., Dodd, T., Farrant, G., and McKernan. A. (1996). Asking the self-employed about their income. Survey Methodology Bulletin 39, 11-15.

Mathiowetz, N., Brown, C., Bound, J. (2002). Measurement error in surveys of the low-income population.. In: Studies of welfare populations: data collection and research issues. Panel on data and methods for measuring the effects of changes in social welfare programs, Ver Ploeg, M., Moffit, R.A., Forbes Citro, C. (eds.). Washington DC, National Academy Press.

Meyer, B.D., and Sullivan, J.X. (2003). Measuring the well-being of the poor using income and consumption. Journal of Human Resources, vol. 38 (5), 1180-1220.

Moore, J.C., Stinson, L.L., and Welniak Jr., E.J. (2000). Income measurement error in surveys: a review. Journal of Official Statistics, vol.16 (4), 331–361.

Micklewright, J., and Schnepf, S.V. (2007). How reliable are income data collected with a single question? IZA DP 3177.

Office for National Statistics. (2010) Harmonized concepts and questions for social data sources. Secondary Sources. Income. Version 3.1.

Rodgers, W., Brown, C., and Duncan, G. (1993). Errors in survey reports of earnings, hours worked, and hourly wages. Journal of the American Statistical Association, vol. 88, 1208-1218.

Romanov, D., and Furman, O. (2011). Imputation and editing of income from the administrative file in the census. UN Economic Commission for Europe, Conference of European Statisticians, Work Session on Statistical Data Editing, Ljubliana.

Romanov, D., and Furman, O. (2006). Analysis of wage data from the 1995 census by using wage file of the National Insurance Institute. Central Bureau of Statistics Working Paper 21, Jerusalem (in Hebrew).

Tourangeau, R., and Smith, T. (1996). Asking sensitive questions: the impact of data collection mode, question format, and question context. Public Opinion Quarterly, vol. 60, 275-304.

Yitzhaki, S. (2007). Pros and cons for using administrative records in statistical bureaus. UN Economic Commission for Europe, Conference of European Statisticians, Geneva.

Yitzhaki, S., and Schechtman, E. (2012). Identifying monotonic and non-monotonic relationships. Economic Letters, vol. 116, 23-25.

Table 1: Distribution of Gross Income in the Social Survey

Band	Income (NIS per month)	Respondents (N)	Pct. respondents in sample
1	Up to 2,000	381	11.2
2	2001–3000	249	7.3
3	3001–4000	336	9.8
4	4001–5000	448	13.1
5	5001–6000	378	11.1
6	6001–7500	378	11.1
7	7501–10000	453	13.3
8	10001–14000	340	10.0
9	14001–21000	233	6.8
10	21000+	221	6.5
Total		3,417	100.0

Table 2: Response to Gross Income Question—Audit Trail Variables

Variable	Employees		Self-employed	
	Mean	S.D.	Mean	S.D.
Time spent responding to item (seconds)	21.8	17.6	27.7	22.3
Frequency of going back to correct gross income after net income question	0.23	0.42	0.22	0.42
Size of correction to response among respondents who made such correction (number of income bands)	0.50	1.69	0.49	2.27

Table 3: Distribution of Measurement Errors

Respondent group	Pct. of negative errors	Pct. of no error	Pct. of positive errors	Mean error	S.D. of error
Full sample	42.1	39.6	18.3	-19.5	71.9
Employees	44.6	40.4	15.0	-21.8	67.3
Self-employed	24.0	33.8	42.2	-2.3	97.7

Table 4: Estimation Results, Full Sample

	All observations	Positive error	Negative error
Intercept	-8.2060	151.1538***	-36.4622
(Ln of) income from administrative file	-4.9720***	-17.0621***	-23.5513***
Correction of response about gross income during interview	-6.8239**	-1.1902	-19.3059***
Size of response correction	4.3593***	0.6398	7.0857**
Response time	0.1343**	0.0292	0.1499
Male	10.9170***	7.3518***	10.8015**
Age	1.1148	0.7864**	7.2169***
Age squared	-0.0147*	-0.0076**	-0.0807***
Married	-5.1054*	-0.3021	0.5351
Arab	-0.9088	-6.4024***	-3.3493
Immigrant from FSU	8.3218**	-1.2411	14.6448**
Education—high school with matriculation	-6.2989*	2.7252	-10.0501
Academic education, B.A.	0.6795	5.6599***	6.2946
Academic education, M.A.	-0.4130	10.5929***	-6.4971
Academic education, Ph.D.	6.8689	19.9953***	18.5741
Status at work: employee	-14.8799***	-17.5435***	31.3204***
Number of jobs held	10.4246***	4.7147***	1.7237
Fear of losing job	-2.7625	-0.5288	-7.4681**
Satisfied with job	-3.3006*	1.6231*	-3.5023
Satisfied with labor income	7.6988***	2.4562***	7.4848**
Household's standard of living has risen	6.7773**	1.6914	7.0980
N	3097	595	1286
Adjusted R-square	0.05	0.49	0.05

Significance level: ***<0.01, **<0.05, *<0.10.

Table 5: Estimation Results, Employees Only

	All observations	Positive error	Negative error
Intercept	-21.9439***	185.0402***	232.9048***
(Ln of) income from administrative file	-6.2019***	-21.0268***	-46.4846***
Correction of response about gross income during interview	-10.1514***	-0.9927	-20.9105***
Size of response correction	2.3582***	0.6322	5.7477***
Male	7.3044**	7.7381***	10.0977*
Arab	-8.9722**	-7.0133***	-14.2879**
Academic education, B.A.	6.4558*	4.7561**	9.5322
Academic education, M.A.	3.7814	15.7729***	-5.1750
Academic education, Ph.D.	9.4299	32.3318***	1.9019
Number of jobs held	8.9031*	4.5159***	3.7518
Usual weekly work hours	5.6238***	0.1757	11.3488***
Satisfied with household's economic situation	7.4893***	1.5024*	8.7456***
Part-time main job	13.8679**	-2.2691	27.9680***
Has worked more hours	0.2307	3.6243**	-1.8295
Has worked fewer hours	8.7946*	-0.6383	10.1158
Received a wage raise	6.5404**	-2.2215	10.7756**
Received a wage cut	-12.3107**	-2.6548	-10.8568
Has not been promoted	-42.9641***	-10.5207*	-67.9578***
Receives full pay for sick days	0.1541	3.4009**	2.7592
Employer participates in pension insurance	-2.6777	-2.4236	21.0496***
Employer participates in advanced-training fund	-16.5640***	6.4326***	-2.6702
Participates in profit-sharing with employer	-3.2737	6.8102**	-6.5944
Has company car	-0.3472	7.6193***	-0.2664
Receives reimbursement of transportation expenses	2.9617	5.4234**	11.1027*
N	2403	344	1107
Adjusted R-square	0.07	0.57	0.09

Significance level: ***<0.01, **<0.05, *<0.10.

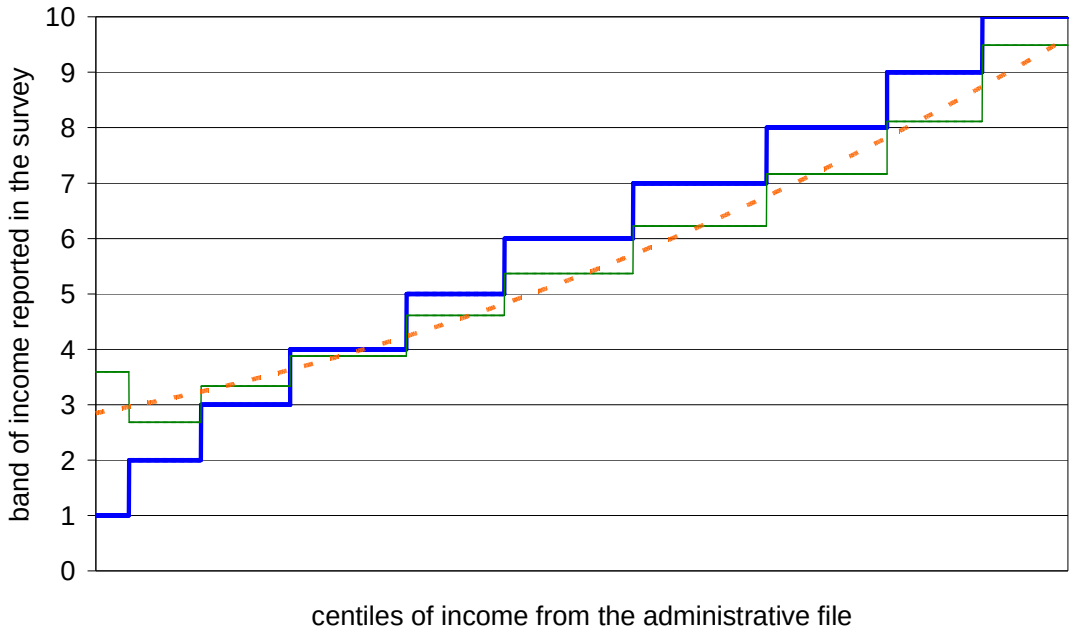
Table 6: Estimation Results, Self-Employed

	All observations	Positive error	Negative error
Intercept	-6.2083	144.0400***	-20.8158
(Ln of) income from administrative file	-4.2012**	-13.6162***	-12.0265
Academic education, B.A.	16.8656	8.1241**	59.9313
Satisfied with job	-7.7512	7.2944***	-32.0741
Satisfied with labor income	23.7797***	1.5338	50.2499***
Household's standard of living has fallen	10.3936	8.5135*	-17.6077
Expects business turnover to increase	-0.3694	-10.5990***	10.4108
Expects business to make progress	-5.9670	-8.9668***	-8.0043
N	369	156	86
Adjusted R-square	0.02	0.34	0.02

Significance level: ***<0.01, **<0.05, *<0.10.

Figure 1: Gross Labor Income

Reportage in Survey vs. Administrative Data

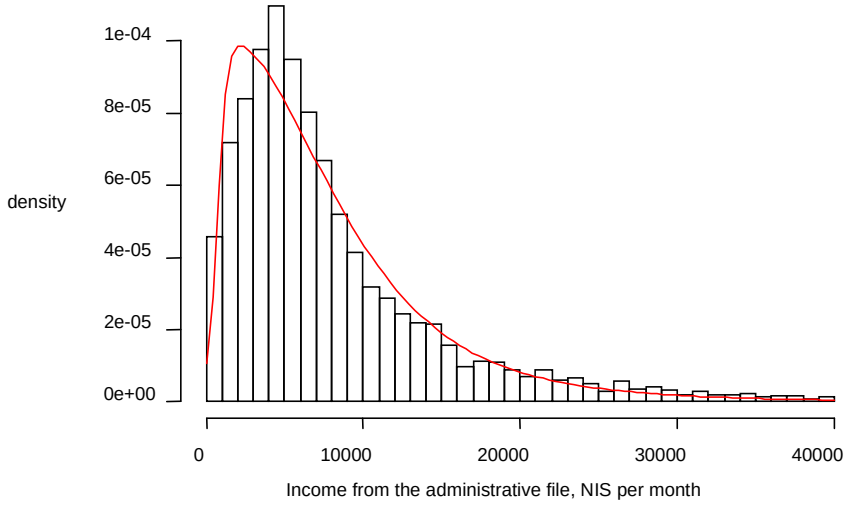


Thick line—income from administrative file by survey reportage bands.

Thin line—income reported in survey.

Broken line—continuous estimator of income reported in survey based on fitted theoretical distribution

Figure 2: Fit of Parametric Distribution
to Respondents' Gross Labor Income Data



Estimated parameters of skew-t distribution: location=698.71, scale=7278.73, shape=13.50, degrees of freedom=3.89.

Figure 3: Measurement Error

as Function of Income from Administrative Source

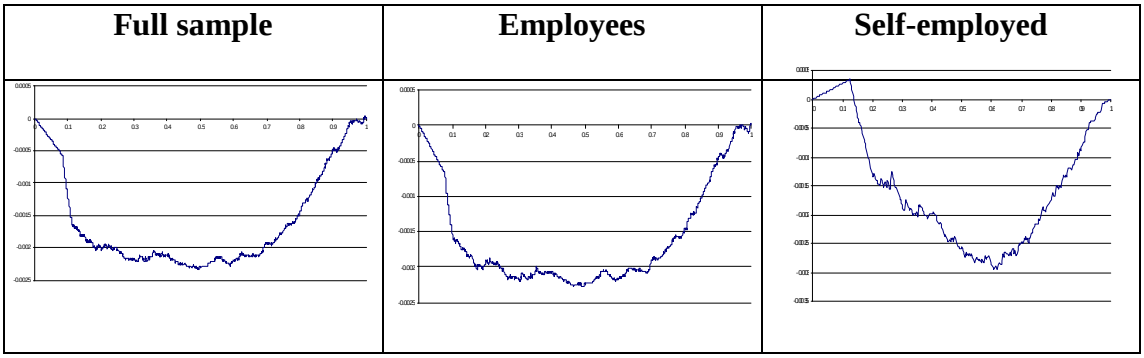


Figure 4: Measurement Error

as Function of Response Time

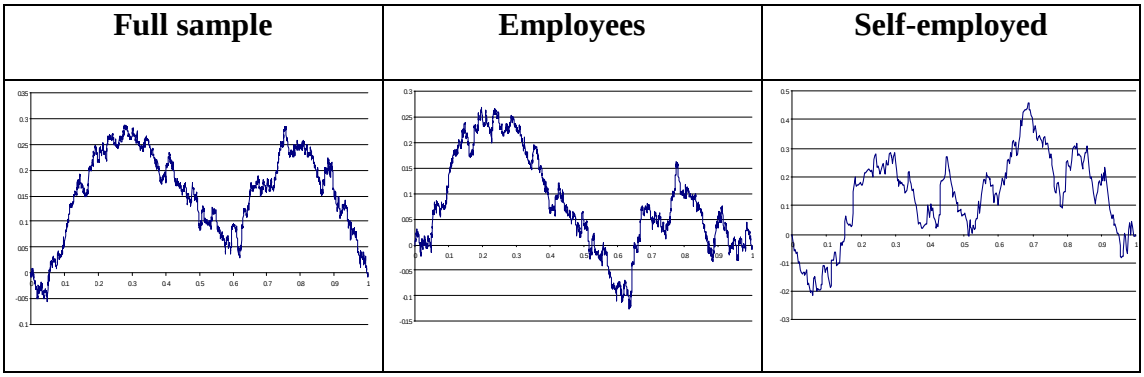
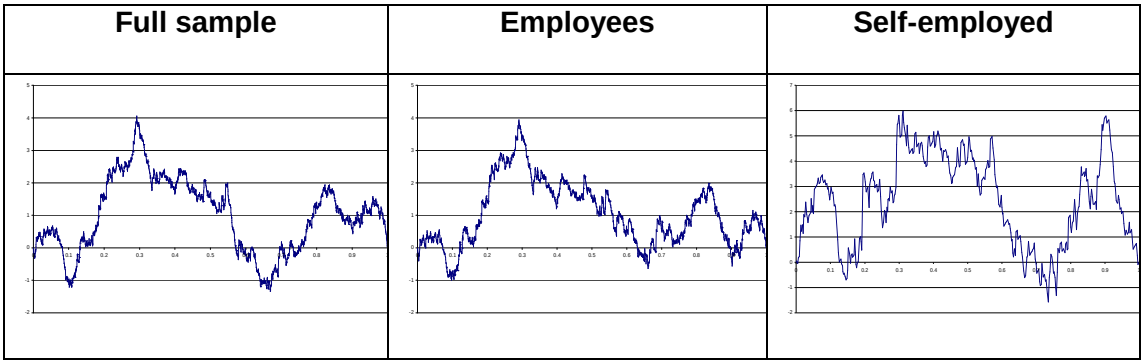


Figure 5: Measurement Error

as Function of Size of Response Correction



Note to Figures 3–5: The horizontal axis indicates the cumulative distribution of the independent variable, X, while the vertical axis presents the cumulative value of the dependent variable, Y. The line of independence is the line connecting (0, 0) with (μy, 1), which may be interpreted as the absolute concentration curve of Y if Y were statistically independent of X.

Appendix: Definition of Variables, Means, and Standard Deviations*

Variable	Employees		Self-employed	
	Mean	S.D.	Mean	S.D.
Measurement error, pct.	-21.82	67.31	-2.29	97.68
Income from administrative file in month preceding survey, NIS	9034.2	9670.2	9649.2	13900.9
Correction of response about gross income during interview (Yes=1, No=0)	0.23	0.42	0.22	0.42
Response time on gross income question, seconds	21.83	17.60	27.73	22.33
Size of response correction, number of income bands as specified in survey	0.50	1.70	0.49	2.27
Male	0.50	0.50	0.72	0.45
Age	39.87	12.42	46.95	13.21
Married	0.67	0.47	0.77	0.42
Arab	0.12	0.32	0.13	0.34
Immigrant from FSU	0.19	0.39	0.07	0.26
Education—high school with matriculation	0.21	0.41	0.15	0.36
Academic education, B.A.	0.19	0.39	0.18	0.38
Academic education, M.A.	0.11	0.31	0.12	0.33
Academic education, Ph.D.	0.01	0.10	0.02	0.14
Number of jobs held	1.10	0.34	1.11	0.37
Usual weekly work hours	42.04	13.88	45.03	18.62
Fear of losing job, subjective evaluation on scale of 1 (no fear) to 4 (acute fear)	1.49	0.74	1.63	0.76
Satisfied with job, subjective evaluation on scale of 1 (totally dissatisfied) to 4 (very satisfied)	3.09	0.77	3.19	0.79
Satisfied with labor income, subjective evaluation on scale of 1 (totally dissatisfied) to 4 (very satisfied)	2.42	0.86	2.53	0.82
Satisfied with household's economic situation, subjective evaluation on scale of 1 (totally dissatisfied) to 4 (very satisfied)	2.49	0.83	2.63	0.78
Household's standard of living has risen in recent years, subjective evaluation (Yes=1, No=0)	0.53	0.50	0.54	0.50
Household's standard of living has fallen in recent years, subjective evaluation (Yes=1, No=0)	0.16	0.37	0.16	0.36
Variables specified for employees only, on basis of self-reportage (Yes=1, No=0)				
Part-time main job	0.06	0.25		
Has worked more hours than usual recently	0.27	0.44		
Has worked fewer hours than usual recently	0.10	0.30		
Received a wage raise on current job	0.59	0.49		

Received a wage cut on current job	0.06	0.24		
Has not been promoted on current job	0.02	0.14		
Receives full pay for sick days	0.56	0.50		
Employer participates in pension insurance	0.75	0.43		
Employer participates in advance-training fund	0.46	0.50		
Participates in profit-sharing with employer	0.11	0.31		
Has company car	0.14	0.35		
Receives reimbursement of transportation expenses	0.17	0.38		
Variables specified for self-employed only				
Expects business turnover to increase			0.32	0.47
Expects business to make progress			0.52	0.50

* The explanatory variables included are those found to have had a significant effect in at least one regression model.